Adapting MPLS Signaling Protocols for the Optical Internet

David W. Griffith
Advanced Network Technologies Division
National Institute of Standards and Technology
100 Bureau Drive, Stop 8920
Gaithersburg, MD 20899-8920

Phone: (301) 975-3512
Fax: (301) 590-0932
Email: david.griffith@nist.gov

Abstract

The Resource ReSerVation Protocol (RSVP) and the Label Distribution Protocol (LDP) were designed by different working groups in the Internet Engineering Task Force (IETF) to accomplish very different tasks. The former was originally developed to support Quality of Service (QoS) for IP networks by allowing end users to reserve bandwidth on intermediate routers. The latter was developed to support distribution of label bindings in networks that support Multi-Protocol Label Switching (MPLS). Both protocols are now being modified to function as the signaling component of an MPLS-based control plane for next-generation optical networks. There has been considerable debate about the relative merits of these two protocols, particularly regarding how they would respond to various failure and restoration scenarios. As a result of this debate, modifications to both protocols to make them more robust have been proposed within the last year. In this article, we will review the operation of the original, "classical" versions of both signaling protocols and we will examine the extensions and modifications that have been proposed to enable them to support agile optical networking.

1 Introduction

In the past decade, the phenomenal growth of the Internet and the attendant rise in demand for bandwidth to support the new services that are being offered on it has driven the massive deployment of a high-speed fiber optic infrastructure. As a result, the large scale carriers are finding that they are now transporting more data traffic than voice traffic, and that the proportion of the total traffic load that is due to data will only increase in the future. Because of this, the requirements for the optical network are moving away from those that resulted in the creation of the ITU-T standards for SONET/SDH in the late 1980s, which were designed primarily to support large scale switched telephone networks. Instead, we are moving toward an optical network architecture that is based on a mesh topology rather than a hierarchical ring topology, and that allows the relatively rapid provisioning of resources to support new connections. In addition, the new optical networking architecture as it is currently envisioned will use control mechanisms that until recently were associated only with router networks.

Optical networks are fundamentally different from the router networks that carry connectionless IP traffic. In its most general form, as described in [18] and [20], the core network can be modeled as a mesh network of Optical Crossconnects (OXCs) that switch traffic from an incoming interface to a corresponding outgoing interface, as shown in Figure 1. Because OXCs establish dedicated paths through a switching fabric, a label mapping does not have the same meaning that it does in an MPLS network as described in [17]. In an optical network, a label mapping that propagates upstream from the network egress carries output (port,wavelength,subchannel) assignments to each OXC, which then assigns a corresponding input (port,wavelength,subchannel) before passing the mapping along.

For over a year, the IETF has been working on two signaling protocols that are being modified to support optical networking, although they were originally designed for other purposes. The Resource ReSerVation Protocol with extensions for Traffic Engineering (RSVP-TE) was originally designed to support Integrated Services over multicast trees in an IP network. It has been extensively modified in recent years to adapt it to other uses, particularly establishing label-switched tunnels in MPLS networks. It continues to evolve as its designers address operational issues (scalability and reliability, for instance) that have previously affected its performance. The Label Distribution Protocol with extensions for Constrained Routing (CR-LDP) was developed specifically to support exchanges of label mapping information in Multiprotocol Label Switching (MPLS) networks. Like RSVP-TE, it is in the process of being extended to support the requirements of optically switched networks.

As these two protocols have evolved, there has been considerable debate within the networking community regarding their relative merits. It is generally agreed that RSVP-TE suffers from scalability problems because it maintains state and exchanges signaling messages on a per-session basis, and because it relies on periodic refreshes to keep its state alive. In addition, RSVP-TE’s use of raw IP or UDP as a transport mechanism has raised concerns regarding its overall reliability. However, basic RSVP as described in [7] is now deployed by a number of IP router manufacturers such as Cisco, Juniper Networks, and 3Com [10]. For this reason, some have argued that RSVP is a logical choice for setting up connections across MPLS and optical networks because it allows interoperation with legacy equipment. CR-LDP, in contrast, uses a different approach to establishing and maintaining connections and does not have some of the problems associated with RSVP because it maintains state on a per-link basis and because it uses TCP as the transport mechanism for sending label mappings and other information.

Several assessments of RSVP-TE versus CR-LDP have already been published that discuss the implementation issues for RSVP that we mentioned above. Since the publication of the most recent of these assessments, the IETF has defined extensions for both protocols to support lightpath establishment across a network of optical cross-connects (OXCs), and there have been additional changes to RSVP-TE.

In this paper we provide a brief overview of the RSVP and CR-LDP protocols with an emphasis on their relative differences. We also summarize the results of the most significant white papers and scholarly articles that compare the two protocols, and we examine extensions to both protocols to support optical networking and also extensions to RSVP-TE to reduce its overhead and improve its reliability.

2 Principal Features of MPLS Signaling Protocols

There are two types of network management, defined in [16], in which we are interested. Configuration management is concerned with the network’s physical plant. It deals with creating and destroying connections and the associated rerouting as well as tracking the status of the equipment. Performance management is just what the name implies; the network operator guarantees a particular level of service to the various users, and monitors compliance by measuring each user's conformance to the traffic characterization that they advertised. Both RSVP-TE and CR-LDP have facilities for supporting both types of network management. In this section, we provide a brief overview of their operation.

2.1 LDP and CR-LDP

The Label Distribution Protocol (LDP) defined in [1] was designed to support the dissemination of label mappings for Forwarding Equivalence Classes (FECs) in MPLS networks. A complete description of the MPLS architecture appears in [17], but we briefly review the essential features here. An illustration of MPLS's operation appears in Figure 2. A FEC is any set of packets that receive the same treatment from an MPLS Label Switching Router (LSR); it is commonly understood to be the set of packets that have a common destination address, but additional criteria for differentiating packets can be used to define FECs. LSRs define Label Switched Paths (LSPs) by binding labels to FECs. Each LSR distributes a label binding by informing the next upstream LSR (*i.e.* the LSR that is next closest to the data source) what label it expects to see attached to all packets that come from the upstream LSR that belong to the FEC. Once label bindings have been distributed, LSRs switch incoming packets by examining the labels that have been attached to them and performing operations on them based on information in the Next Hop Label Forwarding Entry (NHLFE) associated with the relevant labels in the LSR's Label Information Base (LIB). Outgoing packets that have not reached the end of their LSP are given new labels that were bound to the FEC by the LSR immediately downstream. Once a packet reaches the end of its LSP (at the egress LSR), the last label is removed at it continues across the non-MPLS network using whatever forwarding mechanisms are in use there.

LDP relies on Hello messages to establish and maintain LDP adjacencies between label switching routers (LSRs). The protocol allows LSRs that are not physically linked to establish extended adjacencies using Targeted Hello messages, which is useful for creating tunneled label-switched paths (LSPs) across non-MPLS networks or within existing LSPs. CR-LDP Hello messages contain identifiers that allow the LDP peers to keep track of their respective label spaces. LSRs that have at least one Hello adjacency establish LDP sessions using TCP. These sessions transport all CR-LDP signaling information between LSRs. If the LSRs have nothing to send, they maintain the session by periodically transmitting KeepAlive messages. If a KeepAlive message is not received by the time the LDP session timer expires, then the recipient terminates the session and closes the TCP session over which the session was running. In addition, it discards all the labels that were learned in the LDP session. Similarly, if no Hello message is received by the time the Hello interval timer expires, the session terminates and all the labels are lost.

CR-LDP operates using the downstream on demand label distribution mode with ordered control, although LDP itself supports all of the valid MPLS operational schemes defined in [1]. In the downstream-on-demand mode, label bindings are distributed to upstream LSRs only in response to a binding request. If ordered control is being used, an LSR will bind a label to a FEC only if the LSR is the egress LSR for the FEC or if the LSR has received a label binding for the FEC from the next downstream LSR. Thus a new LSP is created when the ingress LSR sends out a Label_Request message, which propagates along the desired path until it reaches the egress LSR. The egress LSR responds to the request by binding a label to the FEC and sending a Label_Mapping message upstream. The Label_Mapping message propagates toward the ingress LSR, establishing LSP state in each intermediate node that it visits. This is shown in Figure 3.

CR-LDP supports LSP setup using strict or loose routes, which can be either free or pinned. CR-LDP carries the full set of traffic parameters for a given LSP in a TRAFFIC TLV in the Label_Request message. These parameters can be modified by intermediate LSRs as the message propagates along the intended path. The set of five parameters can be used to support Intserv QoS, ATM QoS for VBR flows, or QoS for Frame Relay networks. At the egress LSR, the final traffic parameter values are loaded into the Label_Mapping message, which allows all the LSRs on the path to update their reservations as it makes its way back to the ingress. During LSP creation, CR-LDP will send a Notification message to the ingress if any LSR that receives a Label_Request is unable to accommodate the reservation specified in the TRAFFIC TLV.

2.2 RSVP and RSVP-TE

The RSVP protocol, which is described in detail in [7], [24], and [26], was designed to support Integrated Services (IntServ) in IP networks. It does this by reserving resources in routers to achieve a desired quality of service (QoS). RSVP is concerned only with signaling, and it is functionally separate from other networking functions such as routing, admission control, or policy control.

RSVP in its classical form supports receiver-initiated reservations for multicast sessions. Applications with data to transmit, known as senders, advertise their status by transmitting Path messages "downstream" to one or more receivers. As Path messages traverse the network, they establish state information in the RSVP-capable routers that they pass through. This information generally consists of a traffic specification that includes information necessary to support QoS functions (*e.g.* peak data rate, peak burst size, etc.). It also contains information that identify the sender that initiated the Path message and the router immediately upstream (i.e. toward the sender) with respect to the one in question. Once a Path message reaches a receiver, the receiver is able to start sending Resv messages upstream to the entity that sent it. It does not have to do this immediately upon receiving the Path message. Receivers can target more than one sender with their reservation requests; they do this by specifying a FILTERSPEC object that describes the desired sender population. As the Resv message propagates toward the sender, it causes RSVP-capable routers along the route to reserve resources to support the traffic characteristics (or flowspec) that are advertised in the Resv message. When the session's sender receives a Resv message, it can begin sending data. Because the exchange of Path and Resv messages supports only unidirectional flows, a separate set of Path and Resv messages must be exchanged to support a bidirectional session.

RSVP uses soft state to support a flexible response to changes in routing tables or in the characteristics of reservations for individual sessions. Both the Path and Reservation State Blocks for a session will be automatically deleted if they are not periodically refreshed by Path and Resv messages from the senders and receivers. By having the senders and receivers periodically retransmit Path and Resv messages, the network can respond to changes in state (*e.g.* if a receiver drops out of a session) or to changes in routing topology. There is a risk associated with using soft state because RSVP does not run over a transport-layer protocol like TCP. Instead, it runs over either UDP or IP itself (with the protocol number field in the header set to 46). Because RSVP messages can be lost, the retransmission period is typically set to be an integer multiple of the timeout interval. There is a tradeoff between signaling overhead and reliability, *i.e.*, long retransmission interval will reduce the signaling overhead, but it will increase the probability that a reservation will time out due to dropped refresh messages.

Once a receiver is finished receiving data or a sender no longer has data to send, they can delete the state that they created by respectively using ResvTear and PathTear teardown messages. Because RSVP uses soft state, it is possible to avoid using teardown messages and simply allow the state to time out on its own; however, this is discouraged in [7]. Intermediate routers along a path can also send teardown messages if either the Path or the Reservation state associated with a particular session times out. The messages are forwarded immediately by each router that receives them until they reach a node where they do not cause that node's state to change. Because these messages also do not run over a transport layer protocol, they can be lost. In IP networks this is not a fatal problem because the state will time out on its own eventually; when this happens the node whose state times out will transmit the appropriate teardown messages and the ordered teardown process continues.

When the Multiprotocol Label Switching protocol was being developed by the MPLS working group, RSVP was extended to allow it to support requesting and distributing label bindings. These extensions are discussed in [3]. They are used to support the creation of LSP tunnels, i.e., LSPs that are used to tunnel below standard IP-based routing. The modified version of RSVP, known as RSVP-TE, builds support for MPLS into RSVP by defining new objects for transporting label requests and mappings using Path and Resv messages. Details about the FEC that is to receive the mapping are encoded in new versions of the SESSION, SENDER_TEMPLATE, and FILTER_SPEC objects that identify the ingress of the LSP tunnel. In addition, LABEL_REQUEST and LABEL objects have been added to the Path and Resv messages, respectively. RSVP-TE supports only downstream-on-demand label distribution mode with ordered control. The LABEL object allows the Resv message to carry a label stack of arbitrary depth. The LABEL_REQUEST and LABEL objects must be stored in the Path and Reservation state blocks, respectively, of each node in the LSP tunnel, and they must be used in refresh messages, even if there has been no change in the tunnel's state. This will tend to increase the signaling overhead.

If traffic engineering is being used to route LSP tunnels, there are a number of situations where an active tunnel can be rerouted, such as when there is a change of next hop at a particular point in the tunnel or when a node or link failure occurs. When tunnels are rerouted, the preferred course of action is to set up the alternate route before tearing down the existing route so that there is no disruption of the data traffic (this process is known as "make before break"). In order to support this functionality, RSVP-TE must use the shared explicit (SE) reservation style, in which an explicitly-defined set of senders is associated with a given session. This allows the tunnel ingress to specify a tunnel detour associated with a new label descriptor at the same time that it maintains the old tunnel, over which data is still flowing. After the detour has been established, the defunct portion of the tunnel can be torn down or allowed to time out.

RSVP-TE also includes support for explicit routing by incorporating an EXPLICIT_ROUTE object into the Path message and a RECORD_ROUTE object into both the Path and the Resv message. These objects cannot be used in mulitcast sessions. The RECORD_ROUTE object is used to do loop detection, to collect path information and report it to both ends of the tunnel, or to report the path to the tunnel ingress so that it can send an EXPLICIT_ROUTE object in its next Path refresh to pin the route.

RSVP-TE also allows directly connected neighbors to exchange Hello messages to detect node failures. The Hello feature is optional. Hello messages carry either HELLO_REQUEST or HELLO_ACK objects, both of which contain 32 bit instance numbers for the nodes at both ends of the connection. A node that uses the Hello option sends Hello_Request messages to its neighbor at regular intervals (the default interval is 5 msec); a participating recipient replies with a Hello_Ack message. If the sender of the Hello_Request hears nothing from the receiver after a fixed period of time (usually 3.5 Hello intervals), it assumes that communication is lost. If either node resets or experiences a failover, it uses a new instance number in the Hello messages it transmits. This allows an RSVP node to indirectly alert its neighbor that it has reset. If there are multiple numbered links between neighbors (*i.e.* each interface has its own IP address), then Hello messages must be exchanged on all the links.

There have been additional changes to RSVP-TE to address various performance issues and to make it possible to use RSVP as the signaling component of an MPLS-based optical control plane. These enhancements will be discussed in a later section.

3 RSVP-TE and CR-LDP Comparisons

To date, here have been two major studies of RSVP-TE and CR-LDP. Many of the differences between the two protocols have also been summarized online in [23]. Nortel Networks created a white paper [14] whose conclusions also appeared in a more general article on MPLS [11]. At the beginning of this year, Data Connection Limited published a white paper [8] that has attracted considerable attention from the community. In this section we examine the major findings of the Nortel and Data Connection studies, which are summarized in Table 1.

Both studies examine the two protocols with respect to several major criteria, which are the following: reliability (including failure detection), scalability, and interoperability. In addition, the Data Connection study compares the protocols with respect to support for other features such as policy control and traffic control. Both studies conclude that CR-LDP performs better in certain areas (for instance, it clearly scales better than RSVP-TE), although the Data Connection paper offers a more favorable overall assessment of RSVP-TE. In other areas, there are differences between the conclusions of the two assessments. This indicates that there is a need for further study of some of the issues associated with the two protocols.

4 Extensions to RSVP-TE and CR-LDP for Optical Networking

The IETF MPLS Working Group is continuing to make changes to the MPLS architecture; recently the charter has been expanded to support specifying procedures to support header compression across a single link, authentication of LSP originators, and policy support. The group is also working on several other issues, such as supporting optical networking, path protection, and enabling interworking between RSVP-TE and CR-LDP networks. In addition, the IETF RSVP Working Group is continuing to modify RSVP-TE to make it more efficient and reliable, thereby enhancing its utility in larger networks. In this section we describe some of these changes and discuss how they affect the tradeoffs between RSVP-TE and CR-LDP.

4.1 Refresh and Overhead Reduction for RSVP-TE

One of these modifications is the Bundle message, introduced in [5], which encapsulates RSVP messages that are bound for a particular RSVP next hop into the body of a single RSVP message. The Bundle message reduces the number of IP headers that need to be applied to RSVP messages, but it does not reduce the number of messages that need to be transmitted. The size of a Bundle message is limited by the maximum transfer unit (MTU) size of the outgoing link, because it cannot be fragmented. Bundle messages can be used by a node only if the RSVP neighbor that is the intended recipient supports the refresh reduction extensions in [5]. In addition, a node must be careful not to delay state-creating or state-changing Path/Resv messages too much by queueing them until enough have accumulated to form a Bundle message whose size is close to the MTU.

The modifications include new objects to make RSVP more reliable. MESSAGE_IDs can be assigned by RSVP routers to outgoing messages; if the node that sent the message requests an acknowledgement, the recipient will send a MESSAGE_ID_ACK object back to the sender, either piggybacked on another RSVP message or in an Ack message. If the sender of the original message does not receive an acknowledgement within a given amount of time, it will initiate rapid retransmission of the lost message until an acknowledgement is received. This improves reliability but increases overhead, and should be used only for messages that create, modify, or destroy state.

The extensions include a Summary Refresh (Srefresh) message; its purpose is to reduce the amount of bandwidth required to refresh Path and Reservation State in RSVP nodes. It requires that participating nodes use the MESSAGE_ID object extension in Path and Resv messages that are used to create new state or modify existing state. If there is no change to existing state, a Srefresh message containing MESSAGE_ID objects coupled with the IP addresses of the nodes that created the state can be used to carry out the refresh operation, thereby reducing overhead.

4.2 Adaptations for Optical Networking

Any signaling protocol that is used to support setting up optical switched paths (OSPs) must be capable of enabling certain critical functions of optical networks. The protocol must be able to set up bidirectional OSPs whose forward and reverse components typically share the same ports on the intermediate OXCs on the path. It must be able to set up and maintain protection paths, either for individual spans between OXCs or for end-to-end, global repair. It must also be able to set up permanent OSPs that are not taken down unless the entity that created them requests it.

The IETF MPLS Working Group is considering extensions to RSVP-TE and CR-LDP to support these functions; they are in [18] and [20], respectively. Both drafts define an optical label, 8 octets long, that uniquely specifies the port, wavelength, and subchannel that an OXC is using to transport the OSP. A more recent draft explores modifications to RSVP-TE to support optical path setup without resorting to the collision resolution features required in the other two drafts. The operation of these approaches is shown in Figure 4.

4.2.1 Bidirectional OSPs

In [18], the optical label is incorporated into a new LABEL_REQUEST object that is used to set up connections across an inter-OXC span while avoiding collisions. This is done for each connected pair of OXCs in the network by comparing their respective ID numbers. The OXC whose ID has the higher value is the master of the span while the other is the slave. The master is solely responsible for making port assignments. Thus, when the master sends a Path message that contains a LABEL_REQUEST object to the slave, it will tell the slave which interface it should use to support the connection. When a slave sends a LABEL_REQUEST object, it must wait for the master to make a port assignment in a Resv message. To support protection paths, the authors define a new SESSION_ATTRIBUTE object.

A slightly different approach to bidirectional OSP setup using RSVP-TE has been proposed in [13]. Its focus is on reducing setup latency, and so it employs the concept of label suggestion, in which an upstream OXC indicates to its downstream peer the label it would like to have given to it. The benefit of this approach is that the upstream OXC can then begin configuring its switch fabric in anticipation of using the label that it suggested. However, if the downstream OXC is unable to accommodate the suggestion, the upstream OXC will have to abort the preparations it made and reconfigure itself in response to the label that the downstream peer actually assigns to the OSP. This idea is extended to enable RSVP-TE to set up bidirectional OSPs.

The problem of collisions is handled using master/slave designations for each pair of OXCs, as was proposed in [20]. In the approach described in [13], however, the master does not make all label assignments. Rather, the master is the automatic winner in any collision, and it sends a PathErr message to the slave, which then attempts to find an alternative connection across the span. After a fixed number of failed attempts, the slave will give up and send a PathErr message upstream.

A more recent approach to this issue for RSVP-TE has been proposed in [12]; it attempts to avoid some of the complexities associated with the other two approaches. Label assignments for OXCs are made when the OXC receives a Resv message; it then assigns labels for either its two incoming or two outgoing unidirectional interfaces that will support the connection. It informs the downstream OXC of its assignment by it sending a ResvErr message with a new error code; the recipient does not propagate this message. By having each OXC make assignments for both incoming/outgoing interfaces, this approach claims to avoid the contention problems associated with the previous approaches and thus require fewer modifications to RSVP-TE to support contention resolution algorithms. It also adheres to RSVP's original design principle of committing resources to a session only when a Resv message arrives at a node.

4.2.2 Protection Paths

In order to guarantee reliable service and resistance to link and node failures, optical networks provision protection paths. Extensions to both RSVP-TE and CR-LDP have been proposed that allow them to signal what type of protection is requested for an OSP. CR-LDP allows both local and path protection modes to be specified, while RSVP-TE specifies local protection only. The path protection information is opaque to the signaling protocol in both cases.

4.2.3 Permanent OSPs

Permanent OSPs are automatically realized with CR-LDP. To achieve the same result with RSVP-TE, it would be necessary to add a new session attribute, as suggested in [18], that identifies the OSP session as permanent, with the result that the state blocks associated with it never time out. The only way to destroy a session designated as Permanent is with PathTear or ResvTear messages. However, the OSP destruction must be done reliably, *i.e.*, it must occur even if the teardown message is lost. Thus, as stated in [18], either the teardown messages must be made reliable or the OSP must be made non-Permanent through Path updates, so that it will be deleted through state timeout even if the teardown message is lost.

4.3 Interworking Adaptations

Recently the IETF MPLS Working Group began considering the issue of developing an interworking function that would allow LSPs to cross between RSVP-TE and CR-LDP domains. This effort was designed in anticipation of a deployment scenario in which RSVP-TE is deployed at the network edge while CR-LDP is deployed in the network core. To allow the two protocols to communicate in a heterogeneous network, an interworking (IW) LSR is defined in [25] to map signaling messages from one protocol to the other and to handle the different state types (soft versus hard) that exist on opposite sides of the IW LSR. The IW function is complicated by the fact that the mapping between the protocols is not one-to-one.

5 Summary

In this article, we have examined the RSVP-TE and CR-LDP signaling protocols and discussed some of the most significant differences between them. In many cases, CR-LDP attributes make it better suited to the requirements of MPLS or optical networks. However, one of the studies that we have examined asserts that RSVP is more resistant to hardware and software failures because of its reliance on state refresh and IP rather than TCP. This issue requires further study, preferably in the form of quantitative measurements, because it has the potential to generate additional operational requirements for CR-LDP. RSVP-TE is also being modified to make it more efficient and to enable it to operate over optical networks. This, in addition to the work that is being done to make the two protocols interoperate, will maintain both protocols' viability as candidates for signaling in optical networks.

6 References

[1] Andersson, L., *et al.*, "LDP Specification," work in progress, IETF MPLS Working Group, October 1999.
[2] Awduche, D., *et al.*, "Requirements for Traffic Engineering Over MPLS," RFC 2702, September 1999.

[3] Awduche, D., *et al.*, "RSVP-TE: Extensions to RSVP for LSP Tunnels," work in progress, IETF MPLS Working Group, February 2000.
[4] Bala, K., *et al.*, "User Network Interface 1.0 Proposal," OIF2000.125, Optical Internetworking Forum, 23 June 2000.
[5] Berger, L., *et al.*, "RSVP Refresh Overhead Reduction Extensions," work in progress, IETF RSVP Working Group, April 2000.
[6] Bernstein, G., et al., "Optical Domain Service Interconnect (ODSI) Functional Specification," Version 1.3, ODSI Coalition, April 2000.
[7] Braden, R., *et al.*, "Resource ReSerVation Protocol (RSVP): Version 1 Functional Specification," RFC 2205, September 1997.
[8] Brittain, P. and Farrel, A., *MPLS Traffic Engineering: A Choice of Signaling Protocols*, Data Connection Limited, 17 January 2000. (http://www.datacon.co.uk)
[9] Davie, B. and Rekhter, Y., *MPLS: Technology and Applications*, Academic Press, 2000.
[10] Gaines, G. and Fiesta, M., "A Survey of RSVP/QoS Implementations: Update 2," 1 July 1998. (http://www.isi.edu/rsvp/DOCUMENTS/ietf_rsvp-qos_survey_02.txt)
[11] Ghanwani, A., *et al.*, "Traffic Engineering Standards in IP Networks Using MPLS," *IEEE Communications Magazine*, pp. 49-53, December 1999.
[12] Guo, Dan, et al., "Extensions to RSVP-TE for Bi-directional Optical Path Setup," work in progress, IETF Network Working Group, July 2000.
[13] Lang, J., Mitra, K., and Drake, J., "Extensions to RSVP for Optical Networking," work in progress, IETF MPLS Working Group, March 2000.
[14] Li, L., *et al.*, "IP Traffic Engineering Using MPLS Explicit Routing in Carrier Networks: Between the Signaling Approaches—CR-LDP and RSVP," Nortel Networks, April 1999. (http://www.nortelnetworks.com/products/announcements/mpls/collateral/whitepaper.html)
[15] Mathy, L., Hutchison, D., Schmid, S., and Coulson, G., "Improving RSVP for Better Support of Internet Multimedia Communications," *Proceedings of the 1999 6th International Conference on Multimedia Computing and Systems*, vol. 2, pp. 102-106, June 7-11, 1999.
[16] Ramaswami, R. and Sivarajan, K. N., *Optical Networks: A Practical Perspective*, Morgan Kaufmann Publishers, Inc., 1998.
[17] Rosen, E., Viswanathan, A., and Callon, R., "Multiprotocol Label Switching Architecture," work in progress, IETF MPLS Working Group, August 1999.
[18] Saha, D., Rajagopalan, B., and Tang, B., "RSVP Extensions for Signaling Optical Paths," work in progress, IETF MPLS Working Group, March 2000.
[19] Semeria, C., "RSVP Signaling Extensions for MPLS Traffic Engineering," Juniper Networks, Inc., 2000. (http://www.juniper.net/techcenter/techpapers/rsvp_signalling.htm)
[20] Tang, B., Saha, D., and Rajagopalan, B., "Extensions to CR-LDP for Path Establishment in Optical Networks," work in progress, IETF MPLS Working Group, March 2000.
[21] Wang, L., Terzis, A., and Zhang, L., "A New Proposal for RSVP Refreshes," *Proceedings of the 1999 7th International Conference on Network Protocols (ICNP'99)*, pp. 163-172, October 31 – November 3, 1999.
[22] Wang, L., Terzis, A., and Zhang, L., "RSVP Refresh Overhead Reduction by State Compression," work in progress, IETF RSVP Working Group, March 2000.
[23] "A Comparison of Multiprotocol Label Switching (MPLS) Traffic-Engineering Initiatives," Web ProForum Tutorial, The International Engineering Consortium (IEC), (http://www.iec.org/tutorials/mpls_traffic/)
[24] White, P., "RSVP and Integrated Services in the Internet: A Tutorial," *IEEE Communications Magazine*, pp. 100-106, May 1997.
[25] Wright, G., Ballarte, S., and Pearson, T., "CR-LDP Extensions for Interworking with RSVP-TE," work in progress, IETF MPLS Working Group, March 2000.
[26] Zhang, L., *et al.*, "RSVP: A New Resource ReSerVation Protocol," *IEEE Network*, pp. 8-18, September 1993.

Table 1. Comparison of Principal Features of RSVP-TE and CR-LDP.

Issue	Study Results
Reliability	CR-LDP runs over TCP; RSVP-TE runs over UDP or IP. RSVP-TE messages can be lost in transit. RSVP-TE state refresh allows nodes to recover during failover events in which local state is lost. Both use Hello messages, but RSVP-TE Hellos only track whether peer nodes have reset. CR-LDP Hellos identify each peer's label spaces.
Scalability	CR-LDP requires less overhead per session than RSVP-TE, because of RSVP-TE's use of soft state. CR-LDP session maintenance is per-path, not per-LSP. Intermediate LSRs require less state information with CR-LDP [8] if no support for LSP modification. RSVP-TE consumes more CPU cycles because of the refresh function.
Interoperability	Both protocols have undergone interoperability testing. According to [10], as of the summer of 1998, implementations of RSVP as defined in [7] had been tested for interoperability with at least one other implementation by many companies, including 3Com, Bay Networks, Cisco, IBM, and Intel. CR-LDP underwent interoperability testing in November of 1998, involving Nortel Networks, Ericson, and GDC, and Nortel Networks has released the source code for their implementation of CR-LDP. In September of 1999, ITU-T SG13 decided to use CR-LDP to support IP over ATM using MPLS. CR-LDP/RSVP-TE interworking functions have been proposed in IETF [25].
Traffic Control	Both protocols support resource reservation, but RSVP-TE reserves resources on Resv arrival while CR-LDP does it on Label_Request arrival. RSVP-TE Tspec supports IntServ while CR-LDP TRAFFIC TLV supports IntServ as well as ATM and FR QoS. Both support path preemption and have the same setup and holding priority metrics.
Routing	Both protocols support explicit routing, LSP modification, loop detection. RSVP-TE uses RECORD_ROUTE object for loop detection (nothing like Hop Count TLV) while CR-LDP uses Path Vector TLV.

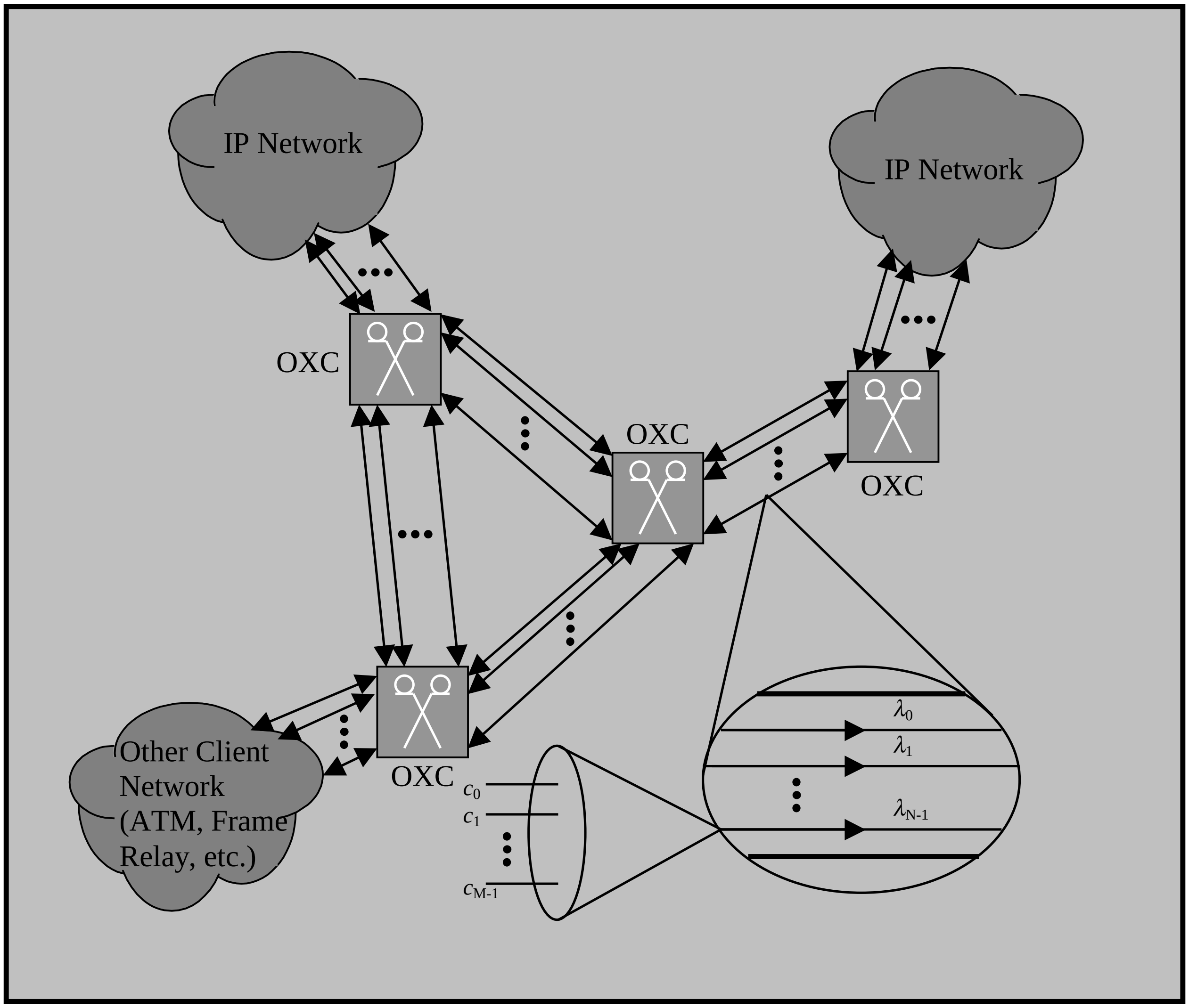

Figure 1. Next-generation optical internet architecture. Client networks supporting various network layer technologies exist at the edge of an optical core that has a mesh topology. Multiple fibers connect each pair of OXCs. In a WDM network, each fiber carries data on multiple wavelengths, which are themselves subdivided into TDM channels. Any set of labels that are distributed to set up a lightpath over such a core network must be able to identify interfaces at sub-wavelength granularities.

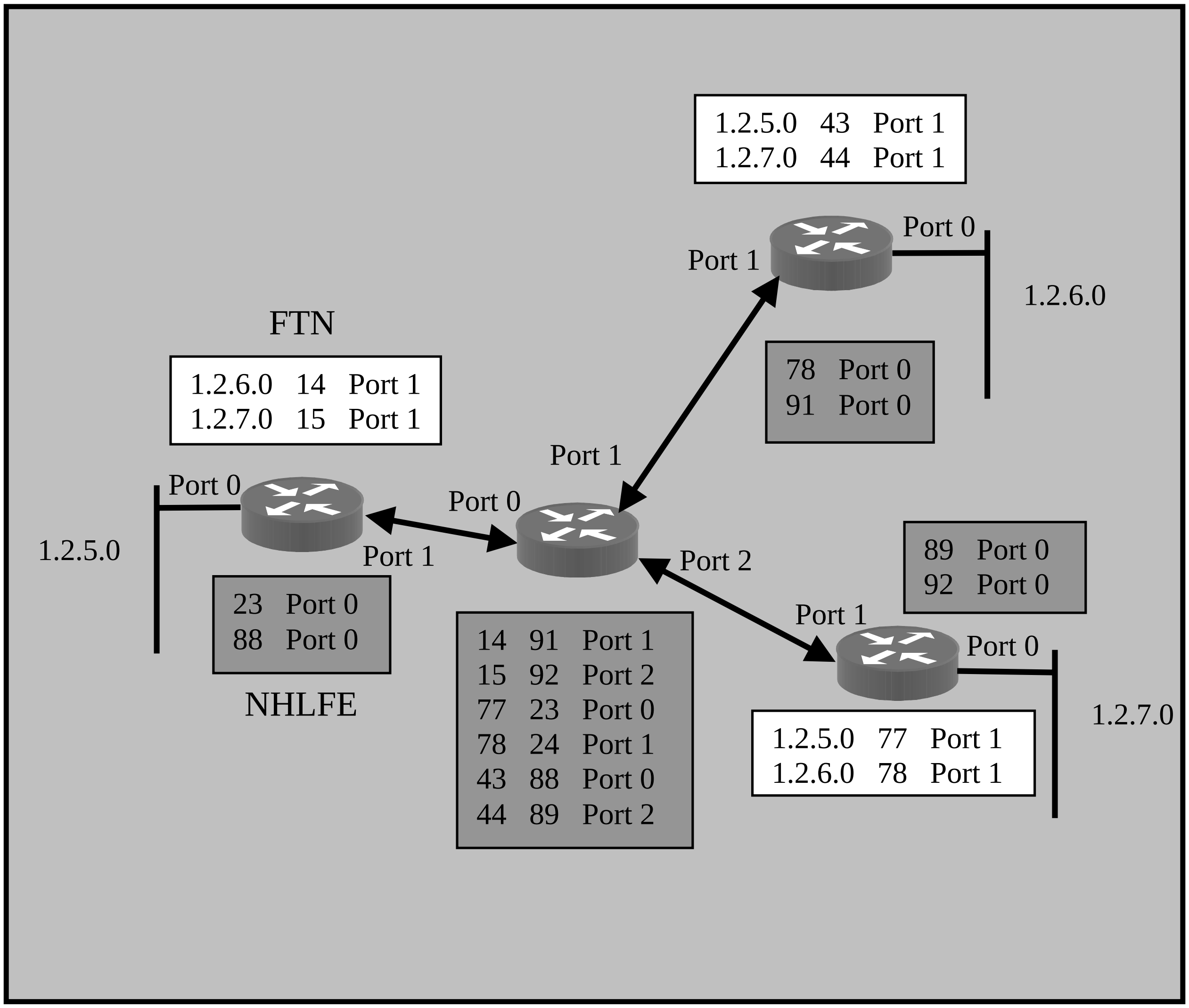

Figure 2. Example of label-based forwarding in an MPLS network. Ingress nodes carry both FEC-to-NHLFE mappings, which map incoming unlabeled packets associated with a given FEC to a Next-Hop Label Forwarding Entry (NHLFE), and tables of NHLFEs. For example, a packet originating on network 1.2.5.0 and destined for network 1.2.7.0 would receive a label of 15 based on the ingress LSR's FTN. The packet would then be forwarded to the router in the center of the network. This router would apply its set of NHLFEs to strip off the packet's label, apply a new label (92), and send the packet out on Port 2. The packet would then arrive at the egress LSR, which would use its own NHLFE to strip off the label and forward the packet on Port 0 to its local Ethernet. Note that it is common practice for the LSR immediately preceding the egress LSR to apply a procedure known as penultimate hop popping, in which the topmost label is stripped off the packet before it is forwarded. The egress LSR will forward the received packet out of the MPLS network without using label switching mechanisms.

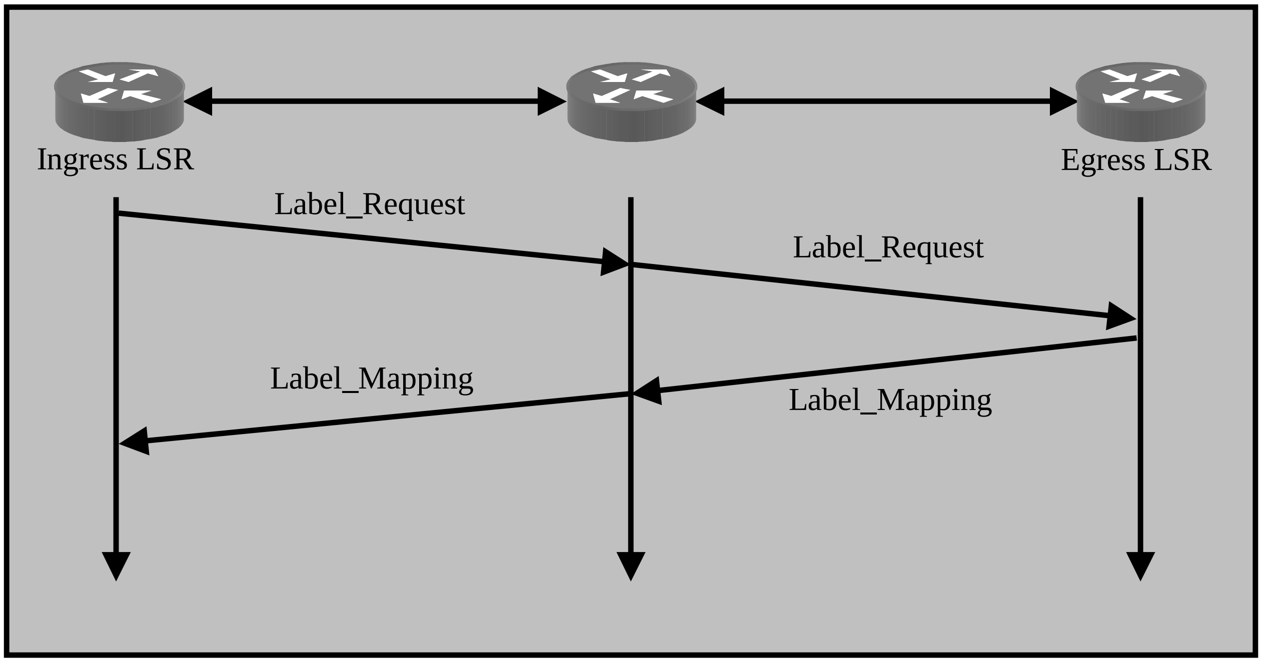

Figure 3. Signal flow in an MPLS network using CR-LDP. The MPLS network must operate using the downstream on demand label distribution mode with ordered control if CR-LDP is being used. A node that has sent a Label_Request message can abort the request by sending a Label_Abort_Request message, although the downstream node will not revoke a label binding that it has made before it received the abort request. Existing label bindings can be destroyed if a node sends a Label_Withdraw message upstream or a Label_Release message downstream.

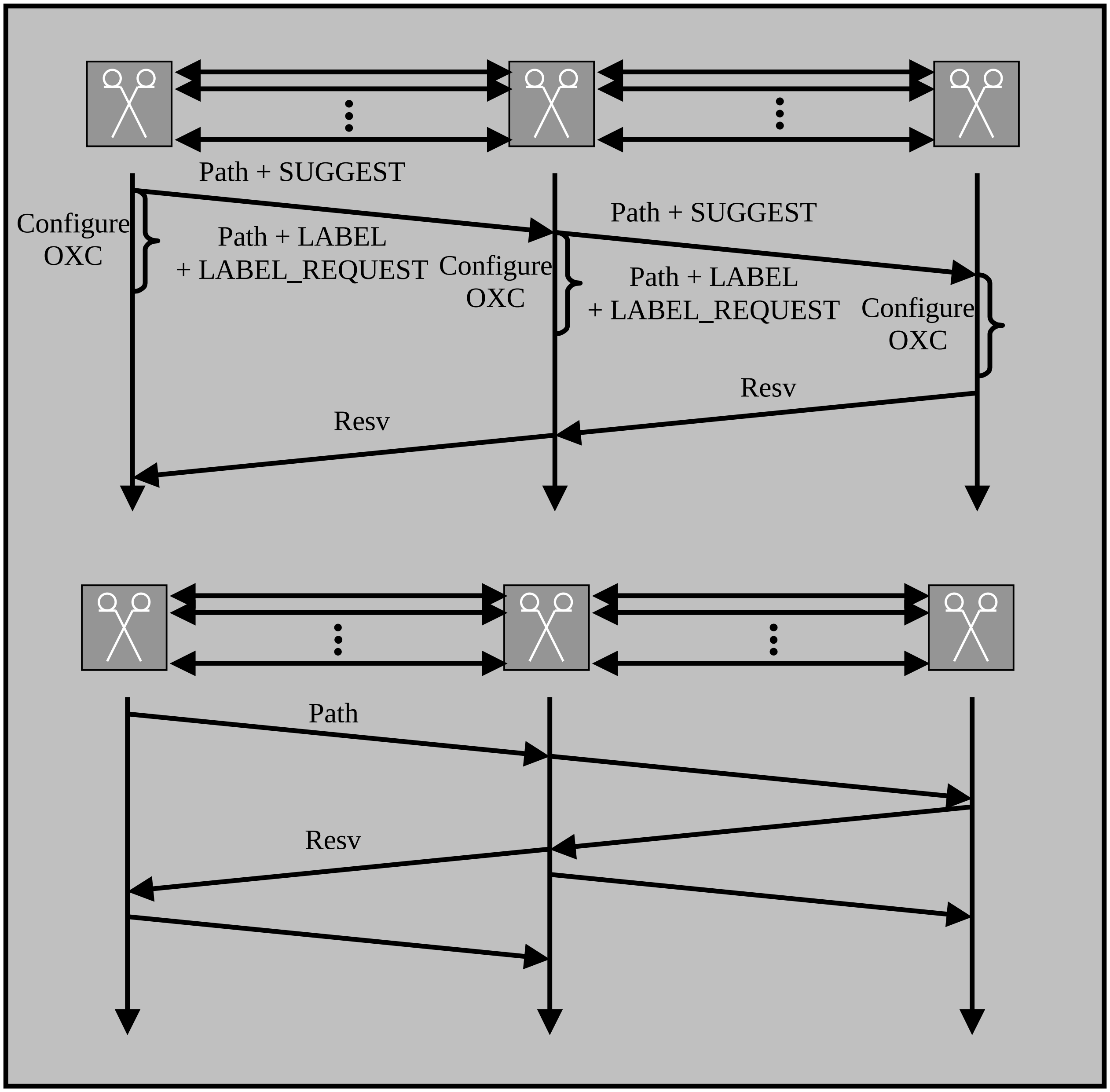

Figure 4. RSVP enhancements to support optical networking. In part (a) we show signal flow associated with the proposals in [13] and [18], which are similar to the proposal for CR-LDP in [20]. In the drafts, labels are suggested by the upstream OXC to its downstream peer in order to reduce the effect of relatively long switch fabric configuration times. In part (b) we show the signal flow associated with the proposal in [12], in which switch configuration does not begin until a Resv message is received.